

RAW SEQUENCE LISTING
ERROR REPORT



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/595,947D

Source: 1600

Date Processed by STIC: 10/27/03

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail. Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom. Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/2003):
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/2003



1600

RAW SEQUENCE LISTING

DATE: 10/27/2003

PATENT APPLICATION: US/09/595,947D

TIME: 09:29:14

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10272003\I595947D.raw

5 <110> APPLICANT: ICARD-LIEPKALNS, Christine
 6 MALLET, Jacques
 7 RAVASSARD, Philippe
 9 <120> TITLE OF INVENTION: POLYPEPTIDES OF THE "BASIC-HELIX-LOOP-HELIX" bHLH
 10 FAMILY, CORRESPONDING NUCLEIC ACID SEQUENCES
 12 <130> FILE REFERENCE: P26,952 USA
 14 <140> CURRENT APPLICATION NUMBER: US 09/595,947D
 15 <141> CURRENT FILING DATE: 2000-06-16
 17 <150> PRIOR APPLICATION NUMBER: FR96/15651
 18 <151> PRIOR FILING DATE: 1996-12-19
 20 <150> PRIOR APPLICATION NUMBER: PCT/FR97/02368
 21 <151> PRIOR FILING DATE: 1997-12-19
 23 <150> PRIOR APPLICATION NUMBER: US 09/331,356
 24 <151> PRIOR FILING DATE: 1999-07-12
 26 <160> NUMBER OF SEQ ID NOS: 40
 28 <170> SOFTWARE: PatentIn Ver. 3.1

pp 1-2, 4-9
Does Not Comply
Corrected Diskette Needed

ERRORED SEQUENCES

129 <210> SEQ ID NO: 5
 130 <211> LENGTH: 18
 131 <212> TYPE: DNA
 132 <213> ORGANISM: Artificial Sequence
 134 <220> FEATURE:
 135 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primers
 137 <400> SEQUENCE: 5 cgcggtgtcc tgcccacc 18
 E--> 137 5 cgcggtgtcc tgcccacc 18 *← move this under <400> 5 line*
 164 <210> SEQ ID NO: 8
 165 <211> LENGTH: 214
 166 <212> TYPE: PRT
 167 <213> ORGANISM: Rattus norvegicus *p.2*
 169 <400> SEQUENCE: 8
 170 Met Ala Pro His Pro Leu Asp Ala Pro Thr Ile Gln Val Ser Gln Glu
 171 1 5 10 15
 173 Thr Gln Gln Pro Phe Pro Gly Ala Ser Asp His Glu Val Leu Ser Ser
 174 20 25 30
 176 Asn Ser Thr Pro Pro Ser Pro Thr Leu Val Pro Arg Asp Cys Ser Glu
 177 35 40 45
 179 Ala Glu Ala Gly Asp Cys Arg Gly Thr Ser Arg Lys Leu Arg Ala Arg
 180 50 55 60
 182 Arg Gly Gly Arg Asn Arg Pro Lys Ser Glu Leu Ala Leu Ser Lys Gln
 183 65 70 75 80

RAW SEQUENCE LISTING

DATE: 10/27/2003

PATENT APPLICATION: US/09/595,947D

TIME: 09:29:14

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10272003\I595947D.raw

```

185 Arg Arg Ser Arg Arg Lys Lys Ala Asn Asp Arg Glu Arg Asn Arg Met
186      85      90      95
E--> 188 His Asn Leu Asn Ser Ala Leu Asp Ala Leu Arg Gly Val Leu Pro Thr
E--> 189      100      105      110
191 Phe Pro Asp Asp Ala Lys Leu Thr Lys Ile Glu Thr Leu Arg Phe Ala
E--> 192      115      120      125
194 His Asn Tyr Ile Trp Ala Leu Thr Gln Thr Leu Arg Ile Ala Asp His
E--> 195      130      135      140
197 Ser Phe Tyr Gly Pro Glu Pro Pro Val Pro Cys Gly Glu Leu Gly Ser
E--> 198 145      150      155      160
200 Pro Gly Gly Gly Ser Ser Gly Asp Trp Gly Ser Ile Tyr Ser Pro Val
E--> 201      165      170      175
203 Ser Gln Ala Gly Ser Leu Ser Pro Thr Ala Ser Leu Glu Glu Phe Pro
E--> 204      180      185      190
206 Gly Leu Gln Val Pro Ser Ser Pro Ser Cys Leu Leu Pro Gly Thr Leu
E--> 207      195      200      205
209 Val Phe Ser Asp Phe Leu
E--> 210      210
266 <210> SEQ ID NO: 10
267 <211> LENGTH: 214
268 <212> TYPE: PRT
269 <213> ORGANISM: Homo sapiens
271 <400> SEQUENCE: 10
272 Met Thr Pro Gln Pro Ser Gly Ala Pro Thr Val Gln Val Thr Arg Glu
273 1 5 10 15
275 Thr Glu Arg Ser Phe Pro Arg Ala Ser Glu Asp Glu Val Thr Cys Pro
276 20 25 30
278 Thr Ser Ala Pro Pro Ser Pro Thr Arg Thr Pro Gly Asn Cys Ala Glu
279 35 40 45
281 Ala Glu Glu Gly Gly Cys Arg Gly Ala Pro Arg Lys Leu Arg Ala Arg
282 50 55 60
284 Arg Gly Gly Arg Ser Arg Pro Lys Ser Glu Leu Ala Leu Ser Lys Gln
285 65 70 75 80
287 Arg Arg Ser Arg Arg Lys Lys Ala Asn Asp Arg Glu Arg Asn Arg Met
288 85 90 95
290 His Asp Leu Asn Ser Ala Leu Asp Ala Leu Arg Gly Val Leu Pro Thr
291 100 105 110
293 Phe Pro Asp Asp Ala Lys Leu Thr Lys Ile Glu Thr Leu Arg Phe Ala
294 115 120 125
296 His Asn Tyr Ile Trp Ala Leu Thr Gln Thr Leu Arg Ile Ala Asp His
297 130 135 140
E--> 299 Ser Leu Tyr Ala Leu Glu Pro Pro Ala Pro His Cys Gly Glu Leu Gly
300 145 150 155 160
302 Ser Pro Gly Gly Pro Pro Gly Asp Trp Gly Ser Leu Tyr Ser Pro Val
303 165 170 175
304 Ser Gln Ala Gly Ser Leu Ser Pro Ala Ala Ser Leu Glu Glu Arg Pro
305 180 185 190
307 Gly Leu Leu Gly Ala Thr Ser Ser Ala Cys Leu Ser Pro Gly Ser Leu
308 195 200 205

```

*see p. 4**for explanation**OK - These numbers are in bold print due to above error**use "L" not "I" numeral "I"*Glu

RAW SEQUENCE LISTING

DATE: 10/27/2003

PATENT APPLICATION: US/09/595,947D

TIME: 09:29:14

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10272003\I595947D.raw

310 Ala Phe Ser Asp Phe Leu
311 210

09/595,947D 4

<210> 8
<211> 214
<212> PRT
<213> Rattus norvegicus

<400> 8

Met Ala Pro His Pro Leu Asp Ala Pro Thr Ile Gln Val Ser Gln Glu
1 5 10 15

Thr Gln Gln Pro Phe Pro Gly Ala Ser Asp His Glu Val Leu Ser Ser
20 25 30

Asn Ser Thr Pro Pro Ser Pro Thr Leu Val Pro Arg Asp Cys Ser Glu
35 40 45

Ala Glu Ala Gly Asp Cys Arg Gly Thr Ser Arg Lys Leu Arg Ala Arg
50 55 60

Arg Gly Gly Arg Asn Arg Pro Lys Ser Glu Leu Ala Leu Ser Lys Gln
65 70 75 80

Arg Arg Ser Arg Arg Lys Lys Ala Asn Asp Arg Glu Arg Asn Arg Met
85 90 95

His Asn Leu Asn Ser Ala Leu Asp Ala Leu Arg Gly Val Leu Pro Thr
100 105 110

↓
What does this mean? Is it a
typographical error? If it represents
a gap, then the amino acids following
it must be in a new sequence ID No.
and the <160> response must be
changed.

09/595,947D 5

<210> 13
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PCR Primer

<400> 13
atcggttgaga ctggtaccag cagagtcacg agagagacta cacggtactg nnnnnnnnnn 60

see p. 6 for env
explanation

6

VARIABLE LOCATION SUMMARY

DATE: 10/27/2003

PATENT APPLICATION: US/09/595,947D

TIME: 09:29:15

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10272003\I595947D.raw

error explanation

Use of n's or Xaa's (NEW RULES):

Use of n's and/or Xaa's have been detected in the Sequence Listing.

Use of <220> to <223> is MANDATORY if n's or Xaa's are present.

in <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

Seq#:2; N Pos. 9,16

Seq#:13; N Pos. 52,53,54,55,56,57,58,59,60

See p.7 for more error

09/595,9420 ?

<210> 26

<211> 15

<212> DNA

<213> Artificial Sequence

<220> ~~<223>~~ Description of Artificial Sequence: Probe

<400> 26

aggaagctcc gggca 15

<210> 27

<211> 1381

<212> RNA

<213> Artificial Sequence

<220> ~~<223>~~ Description of Artificial Sequence: Probe

<400> 27

<2127
no
spaces
within
bracket

insert a hard return

hard return

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/595,947D

DATE: 10/27/2003

TIME: 09:29:15

Input Set : A:\PTO.KD.txt

Output Set: N:\CRF4\10272003\I595947D.raw

L:100 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:2
 L:100 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:2
 L:100 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0
 L:137 M:301 E: (44) No Sequence Data was Shown, SEQ ID:5
 L:137 M:252 E: No. of Seq. differs, <211> LENGTH:Input:18 Found:0 SEQ:5
 L:188 M:333 E: Wrong sequence grouping, Amino acids not in groups!
 L:189 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:8
 M:332 Repeated in SeqNo=8
 L:210 M:252 E: No. of Seq. differs, <211> LENGTH:Input:214 Found:213 SEQ:8
 L:299 M:330 E: (2) Invalid Amino Acid Designator, NUMBER OF INVALID KEYS:1
 L:347 M:258 W: Mandatory Feature missing, <221> Tag not found for SEQ ID#:13
 L:347 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:13
 L:347 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
 L:499 M:256 W: Invalid Numeric Header Field, <220> has non-blank data
 L:501 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:26, <213>
 ORGANISM:Artificial Sequence
 L:501 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:26,Line#:501
 L:510 M:256 W: Invalid Numeric Header Field, <220> has non-blank data
 L:512 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:27, <213>
 ORGANISM:Artificial Sequence
 L:512 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:27,Line#:512